

10 30 50 70  
 gaatlccggcgaglgaggcgctgacagggaactcggggggactctlgacagagccccctggaccacgccgccatcgagacctccag  
 90 110 130 150 170  
 ccagtcctctctgcccgttctctccatggaggccgagccgctccgaggcttcgagcagcgagccggcgaggct  
 190 210 230 250  
 gaccccolgtggcgagggcccggtcctggaggcaggtgcccgtgcccgtcccccgaggcccccgagcgaggctcgggt  
 270 290 310 330  
 cagcatggattcctgggttcatttctgtccctgtttggcagtggtctaatcatgttagtgccacccatgctactacagtttcacc  
 350 370 390 410  
 M D S W F I L V L F G S G L I H V S A N N A T T V S P  
 ttctttgggaacgacagatttaataacatcaacacaggaattggctaaaggaggaataaacctcaaatlcaacctcttc  
 430 450 470 490  
 S L G T T R L I K T S T T E L A K E E N K T S N S T S S  
 agtaatttctcttctlgggcaccacattcagcccaacactgactctggggccacactatgtgactactgttaattcttcaca  
 510 530 550 570 590  
 V I S L S V A P T F S P N L T L E P T Y V T T V N S S H  
 ctctgacaatgggacccaggggcagccagcaggaactcggggcactaccattccccgagcaggaactgacttattgagaa  
 610 630 650 670  
 S D N G T R R A A S T E S G G T T I S P N G S W L I E N  
 ccagltcacggatgccatcaacagaaacccctgggggggaactccagcactgcagcaaccctccagaaactccccccggcaga  
 690 710 730 750  
 Q F T D A I T E P W E G N S S T A A T T P E T F P P A D  
 lgagaccccaattattgagggtgatggccctgtcctctctgctgtaactcgtgttattatcatagttctgtacatgttaag  
 770 790 810 830  
 E T P I I A V M V A L S S L L V I V F I I I V L Y M L R  
 gtttaagaaatcaagcaagctggggagtcatttccaaactcttccgcctgtlcaaatggccgcagcagggatgtggggcccaag  
 F K K Y K Q A G S H S N S F R L S N G R T E D V E P Q S

FIG.1A

FIG. 1B

1610 1630 1650 1670  
 ccctgctggacatgatgcattcggagcgcaaatggatgtataggggttgtagccggatccggcccgctgccagatggt  
 M L D M M H S E R K V D V Y G F V S R I R A Q R C Q M V  
 1690 1710 1730 1750 1770  
 cagacagacatgcagtcgcttcctatataccaggcccttcggagcattatctgtatggggccacagaaactgggaagtgactctc  
 Q T D M Q Y V F I Y] Q A L L E H Y L Y G D T E L E V T S L  
 1790 1810 1830 1850  
 / tagaaaccacccctacaaaaatttatacaagatccacggactccagggactgcacaacccgggttagaggagggttlaagaaatlaacttc  
 E T H L Q K I Y N K I P G T S N N G L E E F K K L T S  
 1870 1890 1910 1930  
 ootcaaaatccagaatgacaagatgcgcacggggaaccltcacagccaacatgaaggaagccgggttllacagatcattccatat  
 I K I Q N D K M R T G N L P A N M K [K N R V L Q I I P Y  
 1950 1970 1990 2010  
 gaatttaacagagtgatcattccagtcacaacggcgaggaaggaacacagactatgtgaacgcattcattgatggataccgc  
 E F N R V I I P V K R G E E N T D Y V N A S F I D C Y R Q  
 2030 2050 2070 2090 2110  
 agaaagactcctacattgccagccaggccctctctccacacgattgaggacttctggcgaatgatctgggagtggaagtcctg  
 K D S Y I A S Q G P L L H T I E D F W R M I W E W K S C  
 2130 2150 2170 2190  
 ttctatcgtaatgctgacagaaactggaaagagagggccagggaggaagtgtgccagtaactggccatctgatggcctgggtgctac  
 S I V M L T E L E E R G Q E K C A Q Y W P S D G L V S Y  
 2210 2230 2250 2270  
 ggagacatcacagttgagctgaaggaagggaggaatgtgaagactacactgtccggagacctctgggtcaccacaccagggaga II  
 G D I T V E L K K E E C E S Y T V R D L L V T N T R E N

FIG.1C

0.1 kb

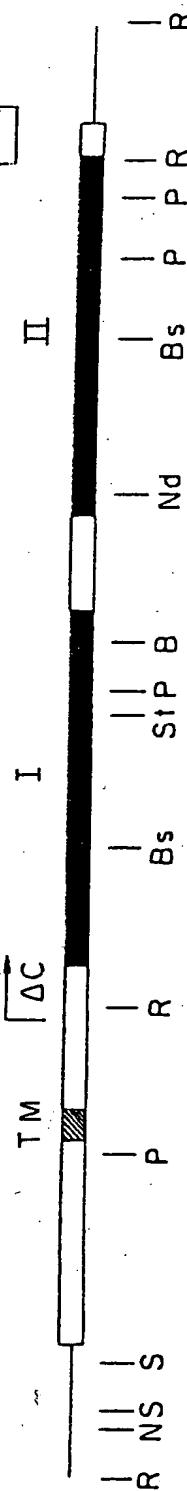


FIG. 1E

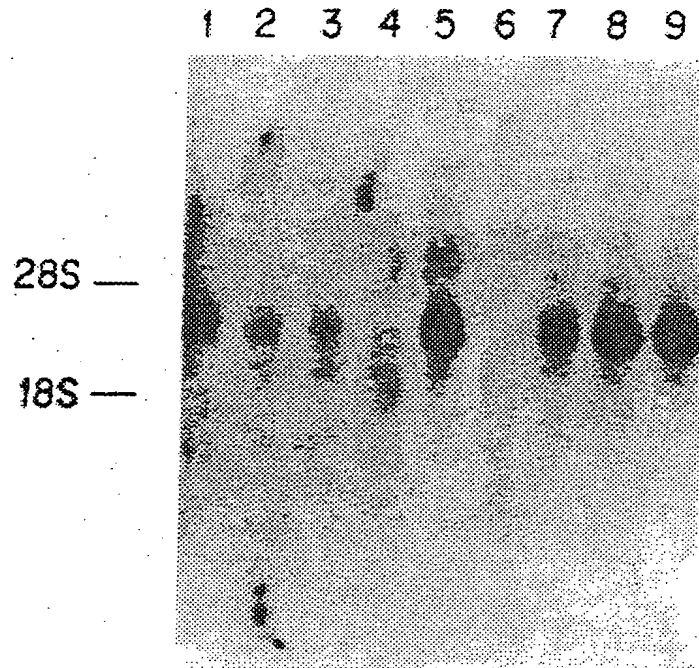


FIG. 2

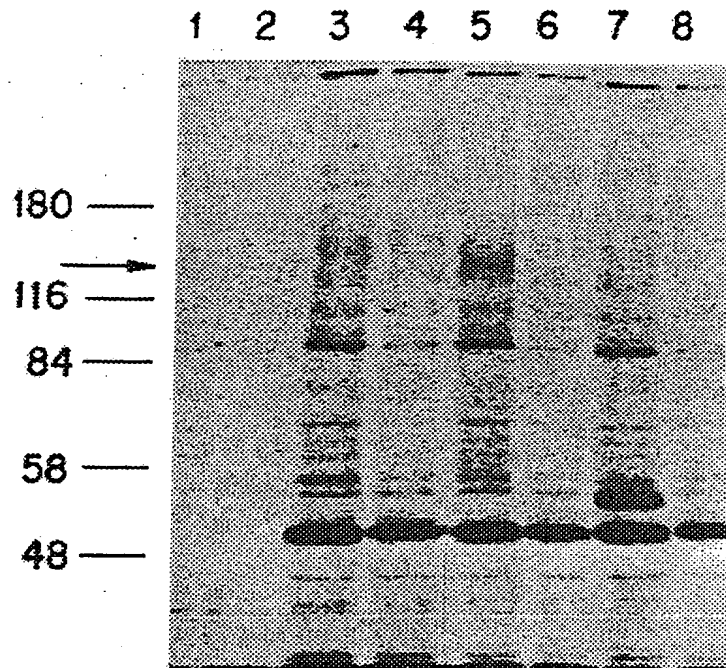


FIG. 3

A vertical scale is shown on the left, ranging from 5' at the bottom to 3' at the top. Major tick marks are labeled 5', 4', 3', and 2'. A hatched rectangular bar is positioned between the 4' and 3' marks, with the label "31-4" centered next to it. To the right of the scale, a horizontal dimension line with vertical end caps is labeled "27-1".

Diagram illustrating the structure of poly(2-vinylpyridine) (PVP) showing the sequence of units I, II, III, IV, V, VI, VII, and VIII. The structure is a linear chain of eight rectangular blocks connected by lines. The blocks are labeled I through VIII from left to right. Block I is a small rectangle, block II is a small rectangle, block III is a small rectangle, block IV is a small rectangle, block V is a large rectangle, block VI is a large rectangle, block VII is a large rectangle, and block VIII is a large rectangle. The chain starts with an NH<sub>2</sub> group on the left and ends with a COOH group on the right.

①  
1 MDSWF ILVLLGSLICVS ANNATTVAPSVGI IRLINSSTAEPVKEEAKTSNPTSSLISLSVAPTFS 68  
2 F H S L T K T T L A N S V I 68  
1 NITLGPTYLT TTVNSSDSNGTTRTASTNSIGITISPNGIWL PDNQFTDARTEPWEGNSSTAATTPETFP 136  
2 L E V H R A E G T S I E I 136  
1 PSGNSDSKDRRDE TPI IAVMVALSSLLVIVF I I I V L Y M I R F K K Y K Q A G S H S N S F R L S N G R T E D V E P Q S 204  
2 A 195  
1 VPLLARSPSTNRKYPPLPVDKLEEE INRRMADDNKLFREEFNALPACPIQATCEAAASKEENKEKNRYV 272  
2 263  
1 NILPYDHSRVHL TPVEGVPSDY INASF INGYQEKKNKF IAAQPKKEETVNDFWRM IWEQNTAT IVMVT 340  
2 (V) 331  
1 NLKERKECKAQYWPDQGCWTYGNIRSVSEDTVLVDYTVRKFCIQQVGDMTNRKPQRL I IQFHFTSW 408  
2 V S V 399  
1 PDFGVPFTPIGMLKFLKKVKACNPQYAGA IWHCSAGVGRGTGTFWIDAMLDMMHTERKVDVYGVFSR 476  
2 S 467  
1 IRAQRQCMVQTDMMQYVF IYQALLEHLYGYDTELEVTSLETHLQIYNKIPGTSNNGLEEEFKLTSIK 544  
2 535  
1 IQNDKMRITGNLPANMKKNRVLQIIPYEFNRVIIPVKRGEENTIDYVNASFIDGYRQKDSYIASQGPLLH 612  
2 (VII) 603  
1 TIEDFWRM IWEWKSCSIVMLTELEERGQEKCAQYWPDSGLVSYGDI TVELKKEECESYTVRDLLVTN 680  
2 671  
1 TRENKSRQIRQFHFGWPEVGIPSDGKGMISIIAAVQKQQQSGNHPITVHCSAGAGRTGTFCALSTV 748  
2 N 739  
1 LERVKAEGILDVFQTVKSLRLQRPHMVQTLQEQEFCYKVVOEYIDAFSDYANFK 802  
2 793

FIG. 4D

	10	20	30	40	
LCA	NqnKNRYVdILPYDynRVeL	sEinGdagSnYINASyldGfkeprKyIAA			
RPTase $\alpha$	NKeKNRYVNILPYDHSRVhLtpvE	GvpdSDYINASfInGYqEknKfIAA			
RPTase $\beta$	NKHKNRYINIVAYDHSRVKLaaqLaekDgKltDYINANYVDGYNrpKAYIAA				
RPTase $\gamma$	NKHKNRYINIIAYDHSRVKLrpLpgKDsKhsDYINANYVDGYNkaKAYIAA				
CON	NkhKNRY-nII-YDhsRVkL—l—k—k—sdYINA—y—dGynepk—yIAa				
	50	60	70	80	90
LCA	QGPrdETVdDFWRMIWEQkatvIVMVTrceEgnrnKCAeYWPsmEegTra				
RPTase $\alpha$	QGPkeETVnDFWRMIWEQntatIVMVTNLkErkeckCAQYWPdqGewTYG				
RPTase $\beta$	QGPLKSTaEDFWRMIWEhNvevIVMITNLVEKGRRKCDQYWPdGSEEEYG				
RPTase $\gamma$	QGPLKSTfEDFWRMIWEqntgiIVMITNLVEKGRRKCDQYWPtenSEEEYG				
CON	QGPik-TveDFWRMIWEqnt-vIVM-TnlvEkgrRK-qYWP—gse-yg				
	100	110	120	130	
LCA	fgdVvVkinqhkrpcDYiIQKI	nIvn	kkekatgRevThiq		
RPTase $\alpha$	NirVsVedVtVLv	DYTVRKFc	IqqvGd	mtnRkpqRIiTQH	
RPTase $\beta$	NfIVTqKSVqVLA	yYTVRnFtIRNTKIKK	Gs	qKGRpsgRVVTQYH	
RPTase $\gamma$	NiIVTIKStkihAc	YTVRrFsIRNTKvKK	GqkgnpKGRqneRVViQYH		
CON	ni-Vlvk-v-vla—dYlvrkf—rntki-k-g-k—kgr—qRvvtqyh				

FIG.5A



	140	150	160	170	180	190
LCA	FTSWPDhGVPedPhlILKlrrrrVnAfsnffsGpIVVHCSAGVGRTGTyigID					
RPTase $\alpha$	FTSWPDfGVPftPi gmlKF l kKVkAcnpqyaGaIVVHCSAGVGRTGTfvVID					
RPTase $\beta$	YTQWPDmGVPEYsLPVL TFVRKaaYAk rhavGPVVHCSAGVGRTGTyIVID					
RPTase $\gamma$	YTQWPDmGVPEYdLPVL TFVRr ssaArmpetGPVIVHCSAGVGRTGTyIViD					
CON	-T-WPDmGVPeypI pVL-fvr-v-aa-----Gp-vVHCSAGVGRTGTyiviD					

	200	210	220	230
LCA	AMLegleaEnKVDVYGyVvkIRrQRCIMVQveaQYiIlhQALvE			
RPTase $\alpha$	AMLdmmhtErKVDVYGfVsrIRaQRCqMVQTdmQYVFlyQALIE			
RPTase $\beta$	SMLQQIqhEgTVNi fGFLKHIRsQRNYLVQTEEQYVF IHDtLvE			
RPTase $\gamma$	SMLQQIkdksTVNvIGFLKHIRtQRNYLVQTEEQYiF IHDaLIE			
CON	-MLqqi-e-V-vyGf-khiR-QR-y-VQteeQY-fIh-aL-E			

FIG. 5B

	10	20	30	40
LCA	NksKNRnsnvIPYdyNRVp	lkhelemskesehdssdsddds	EEpskY	
RPTase $\alpha$	NmkKNRvlqIPYEfNRVilpvkr		GEE <sub>n</sub> TDY	
RPTase $\beta$	NrEKNRLSSIIPvERsRVGIssLs		GE GTDY	
RPTase $\gamma$	NkEKNRnSSvvPsERaRVGlapLp		GmkGTDY	
CON	NkeKNRnss-iPyernRVg	l	geegtdY	
	50	60	70	80
LCA	iNASFImSYwkpevmIAaQGPLkeTlgDFWqMI	fqrKvkvIVMLTELkhg		
RPTase $\alpha$	vNASFIdGYrQkdsyIAaQGPLLHTleDFWRMI	WewKscsIVMLTELeer		
RPTase $\beta$	INASYIMGYYSNEFIIITQHPLLHTIKDFWRMI	WOHNAQIVVMIPDgQnm		
RPTase $\gamma$	INASYIMGYYSNEFIIITQHPLpHTtKDFWRMI	WOHNAQIiVMIPDnQsI		
CON	iNAS-ImGYysnefIIITQ-PLIhTikDFWRMI	wdh-naqiVMI-q		
	100	110	120	130
LCA	dQEiCAQYWgeGkqtYGDleVdLKdtdksstYTl	RvfeIrhskrkdSRTv		
RPTase $\alpha$	gQEiCAQYWPsdGlvSYGDltVeLKkeeeCESYTV	RdlivtntreNkSRqI		
RPTase $\beta$	A EDEFVYWPn	kDEpi	NCESFkVTLmaeehkCLSNEEkil	
RPTase $\gamma$	A EDEFVYWPs	reEsm	NCEaFtVTLiskdrICLSNEEqil	
CON	aE-e-qYWps-g	ygd-v-lk-nces-lvt	e-r-clsne-r-i	
	150	160	170	180
LCA	yQy	qY tnWsvEqIP	aepKellSmIqvVkkQKlpQk	
RPTase $\alpha$	rQf	HF hgWPevgiP	SdgKgmISilaaV Qk Qq	
RPTase $\beta$	IQDFILEATQDDYVLEVRHFQCPKWPNDsPISkTFELISVI			K
RPTase $\gamma$	IhDFILEATQDDYVLEVRHFQCPKWPNDaPISsTFELInVI			K
CON	iqdfileatqddyvlevrhfqcpkwpnpd-Pis-t-ellsvl			qk

FIG.5C

	190	200	210	220	230
LCA	nsseGNkhhkstP	IIiHCrdGsqqTG	iFCALInlLEsaetE	evvDiFQvVKa	
RPTase $\alpha$	qqsGNh	PitVHCsaGagr	TGTFCALsTvLE	rvkaEgiIDVFQt	VKs
RPTase $\beta$	EEAaNR	DGPmIVHDEhGgVt	AGTFCALTTLmh	QLEkENsVDVy	QVAKM
RPTase $\gamma$	EEAItR	DGPtIVHDEyGaVs	AGmICALTTLsq	QLEnENaVDVf	QVAKM
CON	-eea-nr—	dgP—ivH—e—	Gav—GtfCALttll	eqle—En—vDvf	Qv—Km
	240	250			
LCA	LrkaRPgMVstf	EQYqFIYdVias			
RPTase $\alpha$	LaLqRPhMVqT	IEQYefcYKVvqe			
RPTase $\beta$	INLMRPGVF	aDIEQYQFIYKV	iLS		
RPTase $\gamma$	INLMRPGVF	tDIEQYQFIYK	arLS		
CON	-nlmRPg—	iEQYqFIYkv	ils		

FIG.5D

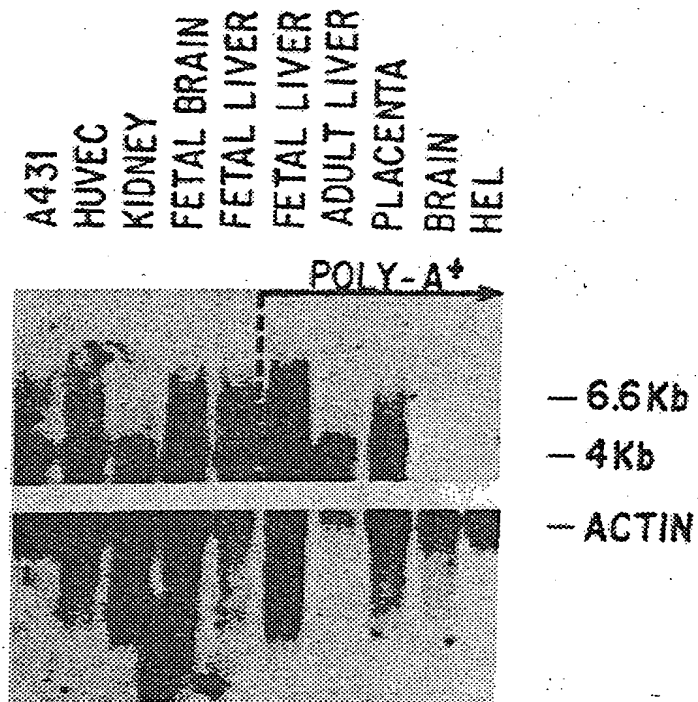
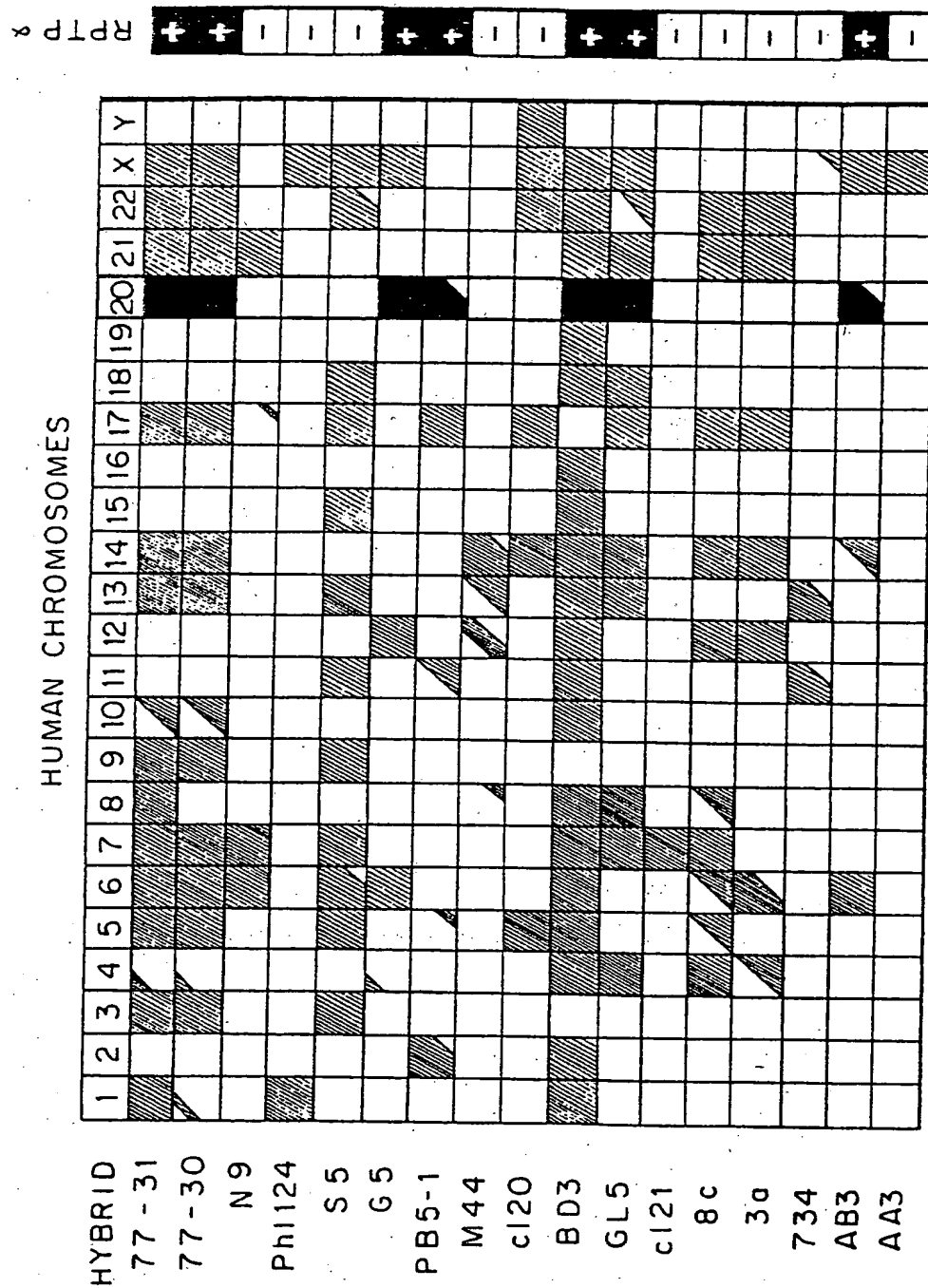


FIG. 6

FIG. 7



1 ATGGATTCTGGTTCATTCTTGTCTGCTCGGCAGTGGTCTGATATGTGTCAGTGCCAAC 60  
1[M D S W F I L V L L G S G L I C V S] A N 20  
SIGNAL PEPTIDE

61 AATGCTACCACAGTTGCACCTTCTGTAGGAATTACAAGATTAACTCATCAACGGCA 120  
21 N A T T V A P S V G I T R L I N S S T A 40

121 GAACCAGTTAAAGAAGAGGCCAAAACTTCAAATCCAACCTTCTTCACTAACTTCTCTTTCT 180  
41 E P V K E E A K T S N P T S S L T S L S 60

181 GTGCCACCAACATTTCAGCCCAAATATAACTCTGGGACCCACCTATTTAACCCTGTCAAT 240  
61 V A P T F S P N I T L G P T Y L T T V N 80

241 TCTTCAGACTCTGACAATGGGACCACAAGAACAGCAAGCACCAATTCTATAGGCATTACA 300  
81 S S D S D N G T T R T A S T N S I G I T 100

301 ATTTACCAAATGGAACGTGGCTTCCAGATAACCAGTTCACGGATGCCAGAACAGAACCC 360  
101 I S P N G T W L P D N Q F T D A R T E P 120

361 TGGGAGGGGAATTCCAGCACCGCAGCAACCACTCCAGAACTTTCCCTCCTTCAGGTAAT 420  
121 W E G N S S T A A T T P E T F P P S G N 140

421 TCTGACTCGAAGGACAGAAGAGATGAGACACCAATTATTGCGGTGATGGTGGCCCTGTCC 480  
141 S D S K D R R D E T [P I I A V M V A L S 160  
TRANSMEMBRANE SEGMENT

481 TCTCTGCTAGTGATCGTGTTTATTATCATAGTTTGTACATGTTAAGGTTTAAGAAATAC 540  
161 S L L V I V F I I I V L Y M L] R F K K Y 180

541 AAGCAAGCTGGGAGCCATTCCAATTCTTTCCGCTTATCCAACGGCCGCACTGAGGATGTG 600  
181 K Q A G S H S N S F R L S N G R T E D V 200

601 GAGCCCCAGAGTGTGCCACTTCTGGCCAGATCCCCAAGCACCAACAGGAAATACCCACCC 660  
201 E P Q S V P L L A R S P S T N R K Y P P 220

661 CTGCCCCGTGACAAGCTGGAAGAGGAAATTAACCGGAGAATGGCAGACGACAATAAGCTC 720  
221 L P V D K L E E E I N R R M A D D N K L 240

721 TTCAGGGAGGAATTCAACGCTCTCCCTGCATGTCCTATCCAGGCCACCTGTGAGGCTGCT 780  
241 F R E E F N A L P A C P I Q A T C E A A 260

781 TCCAAGGAGGAAAACAAGGAAAAAATCGATATGTAAACATCTTGCCTTATGACCACTCT 840  
261 S K E E [N K E K N R Y V N I L P Y D H S 280  
PTPase DOMAIN I

FIG.8A

841 AGAGTCCACCTGACACCGGTTGAAGGGTTCCAGATTCTGATTACATCAATGCTTCATTC 900  
281 R V H L T P V E G V P D S D Y I N A S F 300

901 ATCAACGGTTACCAAGAAAAGAACAAATTCATTGCTGCACAAGGACCAAAAAGAAGAAACC 960  
301 I N G Y Q E K N K F I A A Q G P K E E T 320

961 GTGAATGATTTCTGGCGGATGATCTGGGAACAAAACACAGCCACCATCGTCATCGTTACC 1020  
321 V N D F W R M I W E Q N T A T I V M V T 340

1021 AACCTGAAGGAGAGAAAGGAGTGCAAGTGGCCCCAGTACTGGCCAGACCAAGGCTGCTGG 1080  
341 N L K E R K E C K C A Q Y W P D Q G C W 360

1081 ACCTATGGGAATATTCCGGTGTCTGTAGAGGATGTGACTGTCTTGGTGGACTACACAGTA 1140  
361 T Y G N I R V S V E D V T V L V D Y T V 380

1141 CGGAAGTTCTGCATCCAGCAGGTGGGCGACATGACCAACAGAAAGCCACAGCGCCTCATC 1200  
381 R K F C I Q Q V G D M T N R K P Q R L I 400

1201 ACTCAGTTCCACTTTACCAGCTGGCCAGACTTTGGGGTGCCTTTTACCCGATCGGCATG 1260  
401 T Q F H F T S W P D F G V P F T P I G M 420

1261 CTCAAGTTCTCAAGAAGGTGAAGGCCTGTAACCCTCAGTATGCAGGGGCCATCGTGGTC 1320  
421 L K F L K K V K A C N P Q Y A G A I V V 440

1321 CACTGCAGTGCAGGTGTAGGGCGTACAGGTACCTTTGTCGTCATTGATGCCATGCTGGAC 1380  
441 H C S A G V G R T G T F V V I D A M L D 460

1381 ATGATGCATACAGAACGGAAGGTGGACGTGTATGCCTTTGTGAGCCGGATCCGGGCACAG 1440  
461 M M H T E R K V D V Y G F V S R I R A Q 480

1441 CGCTGCCAGATGGTGCAAACCGATATGCAGTATGTCTTCATATACCAAGCCCTTCTGGAG 1500  
481 R C Q M V Q T D M Q Y V F I Y Q A L L E] 500

1501 CATTATCTCTATGGAGATACAGAACTGGAAGTGACCTCTCTAGAAACCCACCTGCAGAAA 1560  
501 H Y L Y G D T E L E V T S L E T H L Q K 520

1561 ATTTACAACAAAATCCCAGGGACCAGCAACAATGGATTAGAGGAGGAGTTTAAGAAGTTA 1620  
521 I Y N K I P G T S N N G L E E E F K K L 540

FIG.8B

1621 ACATCAATCAAAATCCAGAATGACAAGATGCGGACTGGAAACCTTCCAGCCAACATGAAG 1680  
541 T S I K I Q N D K M R T G N L P A [N M K 560  
PTPase Domain II

1681 AAGAACCGTGTTTTACAGATCATTCCATATGAATTCAACAGAGTGATCATTCCAGTTAAG 1740  
561 K N R V L Q I I P Y E F N R V I I P V K 580

1741 CGGGGCGAAGAGAATACAGACTATGTGAACGCATCCTTTATTGATGGCTACCGGCAGAAG 1800  
581 R G E E N T D Y V N A S F I D G Y R Q K 600

1801 GACTCCTATATCGCCAGCCAGGGCCCTCTTCTCCACACAATTGAGGACTTCTGGCGAATG 1860  
601 D S Y I A S Q G P L L H T I E D F W R M 620

1861 ATCTGGGAGTGGAATCCTGCTCTATCGTGATGCTAACAGAACTGGAGGAGAGAGGCCAG 1920  
621 I W E W K S C S I V M L T E L E E R G Q 640

1921 GAGAAGTGTGCCCAGTACTGGCCATCTGATGGACTGGTGTCTATGGAGATATTACAGTG 1980  
641 E K C A Q Y W P S D G L V S Y G D I T V 660

1981 GAACTGAAGAAGGAGGAGGAATGTGAGAGCTACACCGTCCGAGACCTCCTGGTCACCAAC 2040  
661 E L K K E E E C E S Y T V R D L L V T N 680

2041 ACCAGGGAGAATAAGAGCCGGCAGATCCGGCAGTTCCACTTCCATGGCTGGCCTGAAGTG 2100  
681 T R E N K S R Q I R Q F H F H G W P E V 700

2101 GGCATCCCCAGTGACGGAAGGGCATGATCAGCATCATCGCCGCCGTGCAGAAGCAGCAG 2160  
701 G I P S D G K G M I S I I A A V Q K Q Q 720

2161 CAGCAGTCAGGGAACCAACCCCATCACCGTGCAGTGCAGCGCCGGGGCAGGAAGGACGGGG 2220  
721 Q Q S G N H P I T V H C S A G A G R T G 740

2221 ACCTTCTGTGCCCTGAGCACCGTCCTGGAGCGTGTGAAAGCAGAGGGGATTTTGGATGTC 2280  
741 T F C A L S T V L E R V K A E G I L D V 760

2281 TTCCAGACTGTCAAGAGCCTGCGGCTACAGAGGCCACACATGGTCCAGACACTGGAACAG 2340  
761 F Q T V K S L R L Q R P H M V Q T L E Q 780

2341 TATGAGTTCTGCTACAAGGTGGTGCAGGAGTATATTGATGCATTCTCAGATTATGCCAAC 2400  
781 Y E F C Y K V V Q E] Y I D A F S D Y A N 800

2401 TTCAAGTAA 2409  
801 F K \* 803

FIG.8C